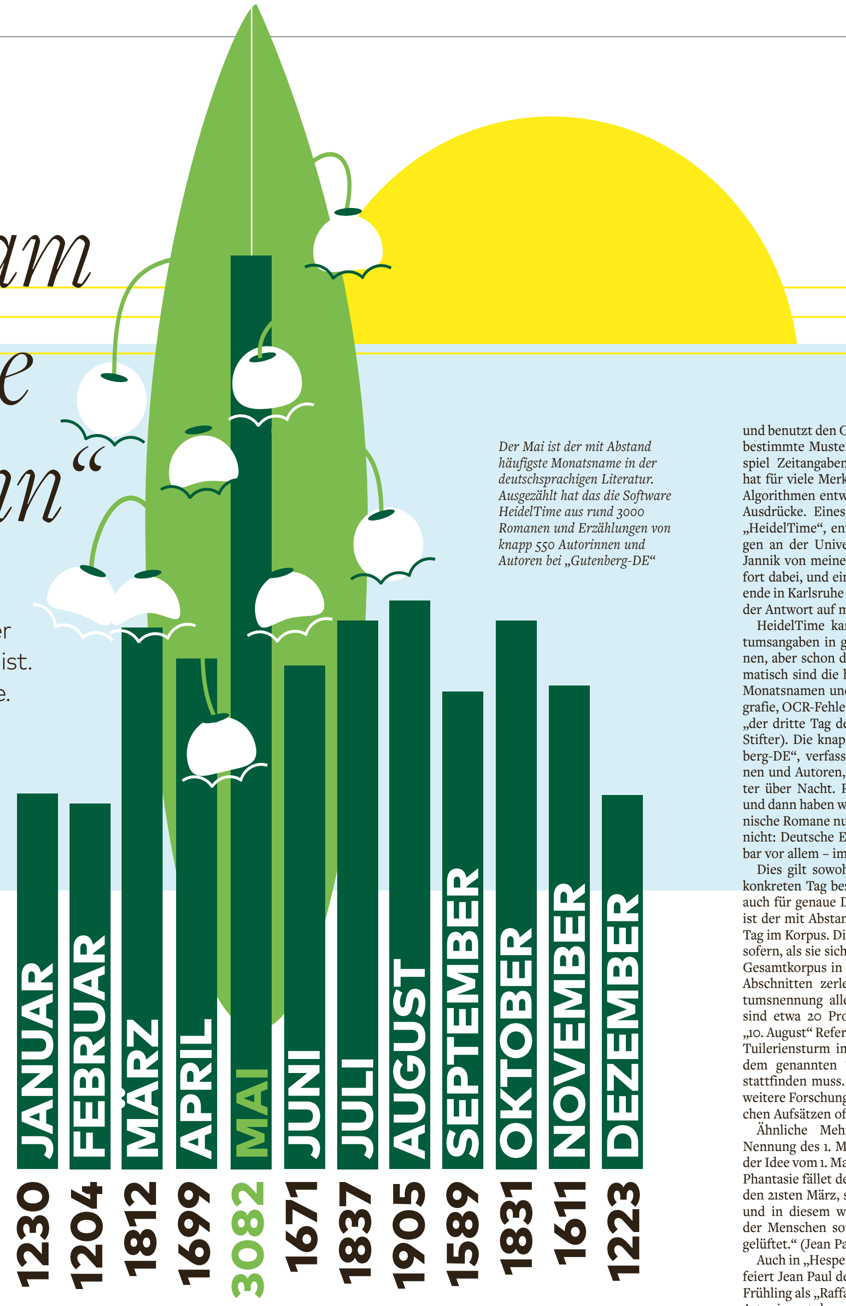
DIE LITERARISCHE WELT
Ein Journal für das literarische Geschehen
Gegründet von Willy Haas, 1925

„Als einst am 1. Mai die Welt begann“

Warum der Wonnemonat in der deutschen Literatur überpräsent ist. Eine datenphilologische Analyse.
Von Frank Fischer



Der Mai ist der mit Abstand häufigste Monatsname in der deutschsprachigen Literatur. Ausgezählt hat das die Software HeidelTime aus rund 3000 Romanen und Erzählungen von knapp 550 Autorinnen und Autoren bei „Gutenberg-DE“

Da ist er wieder, der Monat Mai, der „Mozart des Kalenders“, um ein Gedicht Erich Kästners zu zitieren. Fängt man einmal mit Jahreszeitenlyrik an, ist es schwer, ein Ende zu finden. Goethes „Mailied“: „Wie herrlich leuchtet / Mir die Natur! / Wie glänzt die Sonne! / Wie lacht die Flur!“ Oder Heines „Im wunderschönen Monat Mai“, vertont von Robert Schumann.

Oder Emanuel Geibels „Der Mai ist gekommen, die Bäume schlagen aus“. Oder Friedrich Hagedorns „Der erste Tag im Monat Mai / Ist mir der glücklichste von allen“. Oder das von Mozart vertonte „Komm, lieber Mai, und mache“. Und auch die bitter-selige Zeile „der Mai war mir gewogen“ des Griechen-Müllers, als Gegenbild zu seiner bevorstehenden trüben „Winterreise“. Sicher gibt es auch Gedichte zu anderen Jahresabschnitten („Im traurigen Monat November war's / Die Tage wurden trüber“). Doch der Mai und die deutschsprachige Lyrik, einfach ein Traumpaar.

Was aber ist mit der deutschen Erzählliteratur? Hat auch sie eine Beziehung zum Mai? Und wenn ja, wie sollte man diese messen? Auf die Idee, genau das zu tun, kam ich am 8. Oktober 2014. An diesem Tag lief auf Deutschlandradio Kultur ein Interview mit dem Autor und Übersetzer Stefan Moster. Thema war das damalige Buchmessen-Gastland Finnland, und plötzlich fiel dieser Satz: „Die meisten finnischen Romane spielen im Sommer.“

Aha? Ist das so? Woher weiß Moster das? Ich schrieb ihm eine Mail. Er nannte mir auch sogleich seinen Informanten – einen weiteren Schriftsteller –, aber, um es kurz zu machen, eine zitierfähige Quelle für diesen Fun Fact war nicht mehr aufzutreiben. Halb so schlimm, denn mich interessierte gar nicht so sehr die Richtigkeit der Behauptung, sondern die Art der Information. Und die Frage schloss sich an: Wann spielen eigentlich deutsche Romane?

Es geht also um den Zeitpunkt der Handlungen, nicht darum, wann diese zu Papier gebracht wurden. Das ist zwar auch interessant, aber da begnügen wir uns erst mal mit Rosa Luxemburg, die 1918 in einem Brief aus dem Breslauer Gefängnis von einer Lektürefrucht berichtete, derzufolge „die hervorragendsten wissenschaftlichen und literarischen Produktionen berühmter Männer in die Monate Januar–Februar fallen“.

Doch wie nun weiter? Glücklicherweise war ich in der Zwischenzeit beruflich im Bereich der Digital Humanities tätig geworden, in dem die Beantwortung genau solcher hochskalierter Fragestellungen im Mittelpunkt steht. Möchte man halbwegs belastbare Ergebnisse, benötigt man zunächst ein repräsentatives Korpus. Deutsche Literatur ist inzwischen zuhauf digitalisiert, unter anderem im „Projekt Gutenberg-DE“, das knapp 3000 längere deutsche Erzähltexte enthält, zumeist Romane, die zwischen dem 16. und 20. Jahrhundert veröffentlicht wurden.

Doch bevor wir anfangen zu extrahieren, zu zählen und zu rechnen, ein kurzer Zwischenstopp. Was bedeutet es überhaupt, wenn wir in einem fiktionalen Text einer Datumsangabe begegnen? Es gibt Literaturwissenschaftlerinnen wie Käte Hamburger, die davon ausgehen, dass unser Weltwissen keine Rolle spielt, wenn wir etwa lesen, dass Tristram Shandy am 5. November 1718 geboren wurde oder dass sich „Ulysses“ am 16. Juni 1904 abspielt.

Das dürfte jedoch der allgemeinen Lektüreerfahrung widersprechen. Wir können uns realweltlich zu diesen Daten verhalten, und so haben denn auch einige Leute vor ein paar Monaten Tristram Shandy zum 300. Geburtstag gratuliert, und der 16. Juni wird als „Bloomsday“ jedes Jahr in Dublin und anderswo feierlich begangen. Nur aus diesen Gründen hat es auch Sinn, Datumsangaben aus fiktionalen Texten herauszuziehen und auf einen nicht fiktionalen Kalender zu projizieren. Wirklich fiktive Datumsangaben beschränken sich auf Tage wie den „35. Mai“ im Titel eines Kinderbuches von Erich Kästner („Der 35. Mai oder Konrad reitet in die Südsee“) oder die Nennung eines „80. April“ in Shakespeares „Wintermärchen“. Dies sind aber absolute Ausnahmen, in den Literaturen aller Länder und Zeiten überwiegen die nicht fiktiven Datumsangaben. Zu fünf gewichtigen Beispielen aus der deutschsprachigen Literatur zählen etwa die folgenden:

Mit dem Satz „Den 20. Jänner ging Lenz durchs Gebirg“ beginnt Georg Büchners Erzählung „Lenz“, und es war Paul Celan, der dieses Datum in seiner Büchner-Preis-Rede mit dem Datum der Wannseekonferenz verschränkte und überlegte, ob nicht „jedem Gedicht sein ‚20. Jänner' eingeschrieben bleibt“.

Arthur Schnitzlers „Leutnant Gustl“ sitzt in der Nacht des 4. April nach einem ereignisreichen Tag auf einer Bank im Prater und spielt noch ein wenig mit Selbstmordgedanken.

Für Stefan Zweigs „Dr. B.“ in der „Schachnovelle“ dient eine geschriebene Datumsangabe nachgerade als intellektueller Rettungsanker. Der von allem Geistigen ferngehaltene Inhaftierte „frisst“ beim Anblick eines Kalenders, „in meinem Hunger nach Gedrucktem, nach Geschriebenem diese eine Zahl, diese wenigen Worte ‚27. Juli' an der Wand gleichsam in mein Gehirn hinein“.

Im Fontanejahr 2019 darf auch „Effi Briest“ nicht fehlen, deren Hochzeitstag auf den 3. Oktober fällt – die Folgen ihrer Vermählung mit Innstetten sind bekannt, und es ist irgendwie eine interessante Pointe, dass wir diesen Tag quasi seit 1990 jedes Jahr stellvertretend als Nationalfeiertag mitfeiern.

Begeben wir uns mit dem letzten Datumsexempel ans Jahresende, zum 21. Dezember, an dem sich der junge Werther mit Alberts Pistole den Garaus gemacht hat. Kein Zufall wahrscheinlich, dass das im Dezember passierte und nicht im Mai.

Kaum eine dieser literarischen Datumsangaben dürfte wirklich fest im kollektiven Bewusstsein verankert sein, und prüfungsrelevant sicher auch nicht, im Gegensatz zur einst legendären Examensfrage an der Universität Leipzig: „Wie heißt der Hund von Effi Briest?“

Einzelbeispiele schön und gut, wie kann man nun aber Aussagen treffen über größere Korpora, die man nicht mal ansatzweise selbst lesen kann? Hier helfen die Praktiken des „Distant Reading“ weiter, das heißt, man tritt einen metaphorischen Schritt zurück und benutzt den Computer, um systematisch bestimmte Muster zu extrahieren, zum Beispiel Zeitangaben. Die Computerlinguistik hat für viele Merkmale von Texten Markup-Algorithmen entwickelt, auch für temporale Ausdrücke. Eines dieser Programme heißt „HeidelTime“, entwickelt von Jannik Strötgen an der Universität Heidelberg. Als ich Jannik von meiner Idee erzählte, war er sofort dabei, und ein durchgehacktes Wochenende in Karlsruhe im Januar 2015 brachte uns der Antwort auf meine Ausgangsfrage näher.

HeidelTime kann viel mehr als nur Datumsangaben in großen Textmengen erkennen, aber schon das ist nicht trivial. Problematisch sind die historische Schreibung der Monatsnamen und die schwankende Orthografie, OCR-Fehler oder Formulierungen wie „der dritte Tag des Monates Hornung“ (bei Stifter). Die knapp 3000 Texte aus „Gutenberg-DE“, verfasst von knapp 550 Autorinnen und Autoren, beschäftigen den Computer über Nacht. Ratter, ratter, kring, kring, und dann haben wir die Antwort. Egal ob finnische Romane nun im Sommer spielen oder nicht: Deutsche Erzählliteratur spielt offenbar vor allem – im Mai.

Dies gilt sowohl für den nicht auf einen konkreten Tag bezogenen Monatsnamen als auch für genaue Datumsangaben. Der 1. Mai ist der mit Abstand am häufigsten genannte Tag im Korpus. Die Ergebnisse sind stabil insofern, als sie sich bestätigen, wenn man das Gesamtkorpus in Subkorpora aus 50-Jahres-Abschnitten zerlegt. Dabei kann eine Datumsnennung alles Mögliche bedeuten. So sind etwa 20 Prozent der Nennungen des „10. August“ Referenzen auf den historischen Tuileriensturm im Jahr 1792, ohne dass an dem genannten Tag auch eine Handlung stattfinden muss. Dies lässt also „Raum für weitere Forschung“, wie es in wissenschaftlichen Aufsätzen oft heißt.

Ähnliche Mehrfachfunktionen hat die Nennung des 1. Mai, die oft auch der Aufruf der Idee vom 1. Mai ist: „Im Kalender unserer Phantasie fället der Frühlingsanfang nicht in den 21sten März, sondern in den ersten Mai; und in diesem werden die Kopulierbänder der Menschen sowohl als der Bäume sanft gelüftet.“ (Jean Paul, „Palingenesien“)

Auch in „Hesperus oder 45 Hundposttage“ feiert Jean Paul den Mai und bezeichnet den Frühling als „Raffael der Norderde“. Auf eine Art erinnert das an Harald Schmidt, der 200 Jahre nach Jean Paul jahrelang im Fernsehen seine frühlingshafte Catchphrase aufgesagt hat: „Heute morgen, um 4 Uhr 11, als ich von den Wiesen zurückkam, wo ich den Tau aufgelesen habe“ – was dazu führte, dass Rainald Goetz den Büchern seiner roten Phase den Reihentitel „Heute morgen“ verpasste.

Georg Weerth überhöht den ersten Tag des Wonnemonds sogar noch weiter, wenn auch ähnlich parodistisch in seinem „Schnapphahnski“, dem ersten deutschen Feuilletonroman, vorabgedruckt 1848/1849 in der „Neuen Rheinischen Zeitung“: „Als einst am 1. Mai die Welt begann – ich glaube nämlich, daß die Welt am 1. Mai ihren Anfang nahm und nicht am 1. Januar, wie man fälschlich vermuten möchte, sintemalen die armen nackten Menschen, da sie nicht mit Stiefeln und Sporen auf die Welt kamen, ja im Januar sofort wieder erfroren wären ...“

Romane erweitern also das traditionell von der Lyrik besetzte Themengebiet Mai um interessante Facetten. Am Ende verschwimmen Literatur und Empfindung. So schildert Heine im letzten Abschnitt seiner „Harzreise“, wie ihm der Mai so zu Kopf steigt, dass er „nicht mehr weiß, wo die Ironie aufhört und der Himmel anfängt“. Und er wird fortgetragen von Frühlingsgefühlen: „Es ist der erste Mai, der lumpigste Ladenschwengel hat heute das Recht, sentimental zu werden, und dem Dichter wolltest du es verwehren?“

Frank Fischer lehrt Digital Humanities an der Higher School of Economics in Moskau.